A New argument for the Likelihood Ratio Measure of Confirmation

Glass, DH., & McCartney, M. (2015). A New argument for the Likelihood Ratio Measure of Confirmation. *Acta Analytica*, *30*(1), 59-65. https://doi.org/10.1007/s12136-014-0228-6

Published in:
Acta Analytica

Publication Status:
Published (in print/issue): 01/03/2015

DOI:
10.1007/s12136-014-0228-6

Document Version
Author Accepted version

This is the authors' final version of the paper which was published in Acta Analytica (2015), 30(1), 59-65. The final publication is available at link.springer.com and can be accessed at: http://dx.doi.org/10.1007/s12136-014-0228-6

A New argument for the Likelihood Ratio Measure of Confirmation

David H. Glass and Mark McCartney[1]

Abstract

This paper presents a new argument for the likelihood ratio measure of confirmation by showing that one of the adequacy criteria used in another argument (Zalabardo 2009) can be replaced by a more plausible and better supported criterion which is a special case of the weak likelihood principle. This new argument is also used to show that the likelihood ratio measure is also to be preferred not to a measure that has recently received support in the literature.

Keywords: confirmation, evidential support, inductive support, likelihood ratio.

1. Introduction

Assessing the support that a piece of evidence provides for a given hypothesis is important in many different fields, including law and medicine as well as science more generally. A probabilistic approach to this topic in terms of measures of *confirmation* (or *evidential support*) has not only received significant attention in the philosophical literature but has also been studied in detail in other disciplines such as cognitive science (see for example Tentori *et al.* 2007) and computer science (see for example Greco *et al.* 2012). The idea is that a measure, denoted *c*, of the confirmation of hypothesis, *H*, by evidence, *E*, should satisfy the following condition:

[1] School of Computing and Mathematics, University of Ulster, Shore Road, Newtownabbey, Co. Antrim, BT37 0QB, United Kingdom. E-mail: dh.glass@ulster.ac.uk

$$c(H,E) = \begin{cases} > 0, & if\ Pr(H|E) > Pr(H) \\ = 0, & if\ Pr(H|E) = Pr(H) \\ < 0, & if\ Pr(H|E) < Pr(H), \end{cases}$$

where *Pr* is a probability distribution. If *c(H,E)* > 0, *E* is said to confirm *H*, whereas if *c(H,E)* < 0, *E* is said to disconfirm *H*.

Various confirmation measures have been proposed which satisfy the above condition but are nevertheless not ordinally equivalent.[2] Some well-known measures include the following:

$$d(H,E) = Pr(H \mid E) - Pr(H) \tag{1}$$

$$r(H,E) = \log\left(\frac{Pr(H \mid E)}{Pr(H)}\right) \tag{2}$$

$$l(H,E) = \log\left(\frac{\Pr(E \mid H)}{\Pr(E \mid \sim H)}\right) \tag{3}$$

$$s(H,E) = \Pr(H \mid E) - \Pr(H \mid \sim E) \tag{4}$$

$$ld(H,E) = \Pr(E \mid H) - \Pr(E \mid \sim H) \tag{5}$$

where the distance measure, *d*, the log-ratio measure, *r*, the log-likelihood ratio measure, *l*, in particular have had a lot of supporters.[3] It is worth noting that taking logarithms in (2) and (3) has no effect on the orderings provided by these measures, but is just to ensure that they satisfy the condition noted above. For simplicity, the ratio and likelihood ratio measures, which do not involve logarithms, will be used in the rest of the paper unless otherwise specified

[2] Two measures c_1 and c_2 are said to be ordinally equivalent if and only if for any evidence-hypothesis pairs (E_1,H_1) and (E_2,H_2), $c_1(H_1,E_1) > c_1(H_2,E_2)$ if and only if $c_2(H_1,E_1) > c_2(H_2,E_2)$ and similarly for '<' and '='.

[3] See Zalabardo (2009), Fitelson (1999) and references therein for further details on these measures. The likelihood ratio, *Pr(E|H)/Pr(E|~H)* is often referred to as the Bayes factor due to its role in Bayesian inference where it is the ratio of the posterior to prior odds.

Given the number of non-equivalent confirmation measures, it is not surprising that various adequacy criteria have been proposed to differentiate between them. In this paper, the focus will be on criteria proposed by Zalabardo (2009) in his argument for the likelihood ratio measure. Section 2 will present his argument and show that while one of his criteria is unobjectionable and follows from a widely accepted criterion, the case for another one of his criteria is much less straightforward. An alternative adequacy criterion is proposed in section 3 and it is then shown that this provides reason not only to prefer the likelihood ratio measure to those mentioned already, but also to another measure that has received support in recent literature.

2. Zalabardo's argument for the likelihood ratio measure

In his defence of the likelihood ratio measure of confirmation, José Zalabardo (2009) starts by drawing upon an argument due to Schlesinger (1995) in order to reject the difference measure, *d*. Schlesinger asks us to consider two scenarios, which we shall call *Original Schlesinger Scenarios*, one where $Pr(H_1) = 1/10^9$ and $Pr(H_1|E_1) = 1/100$ and another where $Pr(H_2) = 0.26$ and $Pr(H_2|E_2) = 0.27$. He claims, and Zalabardo agrees, that intuitively the degree of confirmation should be much greater in the former case than in the latter. This is borne out by the ratio measure which gives $r(H_1,E_1) = 10^7$ and $r(H_2,E_2) \approx 1.038$, but not by the difference measure which gives $d(H_1,E_1) = 0.00999\ldots$ and $d(H_2,E_2) = 0.01$. Note that the likelihood ratio measure also handles this example appropriately since $l(H,E)$ can be expressed as $[Pr(H \mid E) \times Pr(\sim H)] / [Pr(H) \times Pr(\sim H \mid E)]$ and so $l(H_2,E_2) \approx 1.01 \text{ x } 10^7$ and $l(H_2,E_2) \approx 1.053$.

Although Schlesinger's argument does not rule out the likelihood difference measure, *ld*, proposed by Nozick (1981) or the measure *s*, proposed by Christensen (1999), Zalabardo appeals to the following adequacy criterion:

(C1) If $Pr(H \mid E_1) > Pr(H \mid E_2)$, then E_1 confirms H to a higher degree than E_2 does.[4]

This is a very plausible criterion and as Fitelson (2007) comments it 'seems to be accepted by all historical practitioners of confirmation theory'. Although this criterion could be discussed further, it will be accepted for the purposes of this paper, where the attention will focus on Zalabardo's argument for the likelihood ratio measure, l, rather than the ratio measure, r.

Zalabardo proposes the following adequacy criteria for a measure of confirmation of hypothesis H by evidence E:

(C2) If $Pr(E_1 \mid H) = Pr(E_2 \mid H)$ and $Pr(E_1 \mid {\sim}H) < Pr(E_2 \mid {\sim}H)$, then E_1 confirms H to a higher degree than E_2 does.

(C3) If $Pr(E_1 \mid H_1) = Pr(E_2 \mid H_2)$ and $Pr(E_1 \mid {\sim}H_1) < Pr(E_2 \mid {\sim}H_2)$, then E_1 confirms H_1 to a higher degree than E_2 confirms H_2.

Zalabardo uses an example to motivate criterion (C2), but the case can be strengthened further by noting that it is a special case of criterion (C1). To see this, note that if $Pr(E_1 \mid H) = Pr(E_2 \mid H)$ and $Pr(E_1 \mid {\sim}H) < Pr(E_2 \mid {\sim}H)$ then it follows that $Pr(E_1) < Pr(E_2)$ and hence via Bayes' theorem that $Pr(H \mid E_1) > Pr(H \mid E_2)$. Criterion (C2) does not help in the selection of a confirmation measure since it is satisfied by all the measures presented in this paper, but Zalabardo uses (C2) to motivate (C3) by arguing that the same intuition that sanctions (C2) also sanctions (C3).

Criterion (C3) is much less straightforward, however. The difference between criteria (C2) and (C3) is that the former considers different evidence for the same hypothesis, whereas the latter

[4] Actually, Zalabardo's criterion is stronger than (C1). It reads 'E_1 confirms H to a higher degree than E_2 does just in case $Pr(H \mid E_1) > Pr(H \mid E_2)$'. However, the weaker condition expressed in (C1) also suffices to rule out ld and s.

applies the same principle to the case where different evidence is considered for different hypotheses. A potential worry might be that the principle will no longer apply since the prior probabilities of the hypotheses will in general be different. This gives rise to a potential conflict between criterion (C3) and Zalabardo's use of Schlesinger's argument. Neither Schlesinger nor Zalabardo try to formulate a criterion based on this example, but in the Original Schlesinger scenarios the idea seems to be that it is the much larger relative increase in probability that warrants the greater degree of confirmation in one case than in the other. Consider now a pair of scenarios, which for reasons that will become apparent we will call *Modified Schlesinger Scenarios*. These scenarios have been chosen to be very similar to those Zalabardo uses to show that the ratio measure, *r*, does not satisfy criterion (C3). In the first scenario, let $Pr(E_1 \mid H_1) = 19/20$, $Pr(E_1 \mid \sim H_1) = 1/40$ and $Pr(H_1) = 2/3$ and in the second let $Pr(E_2 \mid H_2) = 19/20$, $Pr(E_2 \mid \sim H_2) = 1/30$ and $Pr(H_2) = 1/10$. According to criterion (C3), the confirmation should be greater in the former case than in the latter and, of course, this is borne out by the likelihood ratio measure since $l(H_1,E_1) = 38$ and $l(H_2,E_2) = 28.5$. By contrast, the ratio measure fails to satisfy (C3) since $r(H_1,E_1) \approx 1.48$ and $r(H_2,E_2) = 7.6$ since the posterior probabilities are $Pr(H_1 \mid E_1) =76/77$ and $Pr(H_2 \mid E_2) =19/25$. However, in conflict with criterion (C3), the intuition underlying Schlesinger's argument suggests that the degree of confirmation should be much greater in the latter case since it has a much larger relative increase in probability (an increase from 1/10 to 19/25 compared to an increase from 2/3 to 76/77 in the former case).

Does this mean that acceptable measures of confirmation should *fail* to satisfy criterion (C3)? In particular, is there a problem with the ordering provided by the likelihood ratio measure, *l*, in the Modified Schlesinger Scenarios? The answer to this seems to be 'no'. Note that *l* assigns a high value in the case where the probability of H_1 increases from 2/3 to almost 1. More generally, *l* satisfies the following criterion, known as logicality (see Fitelson 2007):

(C4) If *E* entails *H* (~*H*), the degree of confirmation of *H* by *E* should be maximal (minimal).[5]

This criterion makes sense if confirmation is to be understood as a generalization of logical entailment as is appropriate in the context of inductive logic. Equating '*E* entails *H*' with $Pr(H \mid E) = 1$, then it is reasonable that in certain cases, such as the confirmation of H_1 by E_1 in the Modified Schlesinger Scenarios, the degree of confirmation should be high even though the prior probability was high to start with. Thus, although Schlesinger's argument based on the Original Schlesinger Scenarios is very plausible, care must be taken if it is to be applied more generally and there is no clear reason to think that it can be extended in such a way as to pose a problem for the likelihood ratio measure.

Having said that, it is not clear that (C4) should be adopted as an adequacy criterion for confirmation measures. While it is suitable if confirmation is to be considered as a generalization of entailment, it is not so clear that it must be accepted if confirmation is to be considered in terms of the more general notion of evidential support as discussed at the start of this paper, i.e. quantifying the extent to which the hypothesis is made more probable by the evidence.[6] And the same point applies to (C3). While it is far from clear that measures satisfying (C3) should be rejected, given its tension with Schlesinger's argument a more convincing reason would need to

[5] *l* satisfies (C4) provided division by zero is equated with infinity. To avoid this, the ordinally equivalent measure proposed by Kemeny and Oppenheim (1952) can be used instead. It is given by $k(H,E) = [Pr(E|H)-Pr(E|{\sim}H)]/[Pr(E|H)+Pr(E|{\sim}H)]$.

[6] There may well be various ways to think about confirmation. For example, in the field of data mining it has been argued that measures to quantify the strength of association rules should be confirmation measures, but that they should satisfy some criteria differing from those generally accepted in the philosophy literature (Glass, 2013).

be provided to adopt it as an adequacy criterion. Instead, however, an alternative criterion will be proposed in the following section.

3. A new adequacy criterion for confirmation measures

Consider the following proposed adequacy criterion:

(C5) If $Pr(E \mid H_1) = Pr(E \mid H_2)$ and $Pr(E \mid \sim H_1) < Pr(E \mid \sim H_2)$, then E confirms H_1 to a higher degree than E confirms H_2.

Like criterion (C2) and unlike (C3), this proposal keeps something fixed, in this case the evidence, E. Support for (C5) can be obtained by noting that it is a special case of the weak likelihood principle which can be stated as follows (see Joyce 2008):

(C6) If $Pr(E \mid H_1) \geq Pr(E \mid H_2)$ and $Pr(E \mid \sim H_1) \leq Pr(E \mid \sim H_2)$, with one inequality strict, then E confirms H_1 to a higher degree than E confirms H_2.[7]

Joyce argues that this principle 'must be an integral part of any account of evidential relevance that deserves the title "Bayesian"'. Note also that criterion (C5) is a special case of criterion (C3) in which E_1 and E_2 are the same and as such it is a much weaker claim since one can accept criterion (C5) while rejecting criterion (C3) but not vice versa. Furthermore, there is no obvious tension between criterion (C5) and the Original Schlesinger Scenarios or Modified Schlesinger Scenarios. The reason for this is that these scenarios relate to cases where $Pr(H_1 \mid E_1) / Pr(H_1) \neq Pr(H_2 \mid E_2) / Pr(H_2)$, which cannot arise if $Pr(E \mid H_1) = Pr(E \mid H_2)$. For this reason the ratio measure fails to satisfy criterion (C5) since $r(H_1,E_1) = r(H_2,E_2)$ if $Pr(E \mid H_1) = Pr(E \mid H_2)$ irrespective of the values of $Pr(E \mid \sim H_1)$ and $Pr(E \mid \sim H_2)$. Clearly, the likelihood ratio measure satisfies (C5).

[7] Joyce also states a slightly different version which does not allow for equality between $Pr(E \mid H_1)$ *and* $Pr(E \mid H_2)$. Clearly, (5) is not a special case if it is stated in that way, but it would at most be a very modest extension of it.

It is worth noting that criterion (C5) is not as discriminating as criterion (C3) since, apart from the ratio measure, (C5) is satisfied by all the other measures considered so far. However, if providing a correct ordering for the Original Schlesinger Scenarios and criterion (C1) are both to be imposed as requirements for confirmation measures, as Zalabardo proposes, *d*, *s* and *ld* can all be ruled out. Since (5) rules out *r*, this means that *l* is the only measure satisfying the proposed adequacy requirements.

Another confirmation measure, which has not been considered so far, is the certainty factor (Shortliffe and Buchanan 1975), which has been used in the field of expert systems and has recently been advocated as a measure of confirmation (see Crupi *et al.* 2007 and Crupi and Tentori 2013). It is defined as follows:

$$cf(H,E)=\begin{cases}\dfrac{Pr(H|E)-Pr(H)}{1-Pr(H)}, & if\ Pr(H|E)\geq Pr(H)\\[2ex] \dfrac{Pr(H|E)-Pr(H)}{Pr(H)}, & if\ Pr(H|E)<Pr(H)\end{cases}\tag{6}$$

Both *l* and *cf* also satisfy symmetry criteria proposed by Eells and Fitelson (2002), while *cf* also satisfies extended symmetry criteria proposed by Crupi *et al.* (2007) but *l* does not.[8] The symmetry proposals of Crupi *et al.* are based on generalizing those of classical logic. For example, if *E* entails *H* it is not the case in general that *H* also entails *E* and so, taking the degree of confirmation to be a generalization of entailment, they argue that if *E* confirms *H* the degree of confirmation of *H* by *E* should not necessarily be the same as that of *E* by *H*. However, in

[8] More precisely it is the log-likelihood ratio that satisfies the symmetry requirements of Eells and Fitelson. The measure of Kemeny and Oppenheim (see footnote 4) also satisfies these requirements.

classical logic, if *E* refutes *H* (i.e. *E* entails $\sim H$), then *H* also refutes *E* and so they argue that if *E* disconfirms *H* the degree of disconfirmation of *H* by *E* should be the same as the degree of disconfirmation of *E* by *H*.

A proper evaluation of the interesting proposal of Crupi *et al.* is beyond the scope of this paper, but it is worth noting that although *l* does not satisfy their symmetry principles, it does satisfy them in the extreme cases, for example, when *E* refutes *H*. This follows from the fact that *l* satisfies (C4). Arguably this is sufficient in terms of symmetry for a confirmation measure to generalize entailment in classical logic. At the very least, the failure of *l* to satisfy the more detailed symmetry criteria of Crupi *et al.* does not seem sufficient to rule it out as a measure of confirmation independent of other criteria.

Another factor to take into account when comparing *l* and *cf* is the fact that *cf* does not handle the Original Schlesinger Scenarios in a satisfactory way. In the scenario where $Pr(H_1) = 1/10^9$ and $Pr(H_1 \mid E_1)= 1/100$ the certainty factor gives the result $cf(H_1,E_1) = 0.00999\ldots$ while in the scenario where $Pr(H_2) = 0.26$ and $Pr(H_2 \mid E_2)= 0.27$ it gives $cf(H_2,E_2) \approx 0.0135$. As pointed out earlier, *l* does give an intuitively correct result in this case.

Finally, *cf* does not satisfy criterion (C5) in cases of disconfirmation. To see this, suppose that the antecedent of (C5) holds so that $Pr(E \mid H_1) = Pr(E \mid H_2)$ and $Pr(E \mid \sim H_1) < Pr(E \mid \sim H_2)$. Let us further suppose that *E* disconfirms H_1, which means that $Pr(E \mid H_1) < Pr(E)$ and hence $Pr(E \mid H_2) < Pr(E)$ so that *E* disconfirms H_2 as well. Using Bayes' theorem to replace the term $Pr(H \mid E)$ in the expression for $cf(H,E)$ in the case of disconfirmation, $cf(H,E)$ can be expressed as

$Pr(E|H)/Pr(E)$ – 1 and so $cf(H_1,E)= cf(H_2,E)$ which disagrees with criterion (C5). Strangely, although *cf* fails to satisfy (C5) in cases of disconfirmation, it does satisfy it in cases of confirmation. Hence, according to *cf*, whether the difference between $Pr(E \mid {\sim}H_1)$ and $Pr(E \mid {\sim}H_2)$ results in a difference in degree of confirmation in cases where $Pr(E \mid H_1) = Pr(E \mid H_2)$ depends on whether confirmation or disconfirmation occurs.[9]

Overall, these findings give us reason to prefer the likelihood ratio measure, *l*, to the certainty factor, *cf*, as well as to the other measures presented earlier in the paper. While *cf* does satisfy some symmetry properties not satisfied by *l*, it is far from clear that this constitutes a serious problem for *l*.

5. Conclusion

This paper has provided a new defence of the likelihood ratio measure based on criteria (C1) and (C5) and an argument due to Schlesinger. This is a development of an earlier argument due to Zalabardo (2009). In particular, it has been argued that one of Zalabardo's adequacy criteria (C3) can be replaced by a logically weaker and better supported criterion (C5). It has also been argued that these same criteria give us reasons to prefer the likelihood ratio measure not only to a number of measures that have received a lot of attention in the philosophical literature but also to the certainty factor measure which has been advocated recently.

References

Christensen D. (1999). Measuring confirmation. Journal of Philosophy, 96: 437–61.

[9] Crupi *et al.* (2007) state that *cf* satisfies the weak likelihood principle, so how can it fail to satisfy (C5) which is a special case of it? This is due to slightly different formulations of the principle (see footnote 6).

Crupi, V., Tentori, K. and Gonzalez, M. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. Philosophy of Science, 74: 229–252.

Crupi, V. and Tentori, K. (2013). Confirmation as partial entailment: A representation theorem in inductive logic. Journal of Applied Logic, in press. doi: 10.1016/j.jal.2013.03.002

Eells, E. and Fitelson, B. (2002). Symmetries and asymmetries in evidential support. Philosophical Studies, 107: 129–142.

Fitelson, B. (1999). The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity. Philosophy of Science, 66: S362–S378.

Fitelson, B. (2007). Logical foundations of evidential support. Philosophy of Science, 73: 510–512.

Glass, D.H. (2013). Confirmation measures of association rule interestingness. Knowledge-Based Systems, 44: 65-77.

Greco, S., Słowiński, R. and Szczęch, I. (2012). Properties of rule interestingness measures and alternative approaches to normalization of measures. Information Sciences, 216: 1–16.

Joyce, J. (2008). Bayes' Theorem. In The Stanford Encyclopedia of Philosophy (Fall 2008), ed. E. N. Zalta. URL = http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/

Kemeny, J.G. and Oppenheim, P. (1952). Degree of factual support. Philosophy of Science, 19: 307–324.

Nozick, R. (1981). Philosophical Explanations. Cambridge, Mass.: Harvard University Press.

Schlesinger, G.N. (1995). Measuring degrees of confirmation. Analysis, 55: 208–12.

Shortliffe, E. and Buchanan, B. (1975). A model of inexact reasoning in medicine. Mathematical Biosciences, 23: 351–379.

Tentori, K., Crupi, V., Bonini, N. and Osherson, D. (2007). Comparison of confirmation measures. Cognition, 103: 107–119.

Zalabardo, J. (2009). An argument for the likelihood-ratio measure of confirmation. Analysis, 69: 630-635.